

Section 7: Solutions

Review of Main Concepts

- **Multivariate: Discrete to Continuous:**

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$	$f_{X,Y}(x, y) \neq \mathbb{P}(X = x, Y = y)$
Joint range/support $\Omega_{X,Y}$	$\{(x, y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x, y) > 0\}$	$\{(x, y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x, y) > 0\}$
Joint CDF	$F_{X,Y}(x, y) = \sum_{t \leq x, s \leq y} p_{X,Y}(t, s)$	$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$
Normalization	$\sum_{x,y} p_{X,Y}(x, y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x, y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$
Expectation	$\mathbb{E}[g(X, Y)] = \sum_{x,y} g(x, y) p_{X,Y}(x, y)$	$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$
Independence must have	$\forall x, y, p_{X,Y}(x, y) = p_X(x) p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x, y, f_{X,Y}(x, y) = f_X(x) f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

- **Law of Total Probability (r.v. version):** If X is a discrete random variable, then

$$\mathbb{P}(A) = \sum_{x \in \Omega_X} \mathbb{P}(A|X = x) p_X(x) \quad \text{discrete } X$$

- **Law of Total Expectation (Event Version):** Let X be a discrete random variable, and let events $A_1, \dots, A_n$ partition the sample space. Then,

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X|A_i] \mathbb{P}(A_i)$$

- **Conditional Expectation:** See table. Note that linearity of expectation still applies to conditional expectation:
 $\mathbb{E}[X + Y|A] = \mathbb{E}[X|A] + \mathbb{E}[Y|A]$

- **Law of Total Expectation (RV Version):** Suppose X and Y are random variables. Then,

$$\mathbb{E}[X] = \sum_y \mathbb{E}[X|Y = y] p_Y(y) \quad \text{discrete version.}$$

- **Conditional distributions**

	Discrete	Continuous
Conditional PMF/PDF	$p_{X Y}(x y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X Y}(x y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$\mathbb{E}[X Y = y] = \sum_x x p_{X Y}(x y)$	$\mathbb{E}[X Y = y] = \int_{-\infty}^{\infty} x f_{X Y}(x y) dx$

- **Continuous Law of Total Probability:**

$$\mathbb{P}(A) = \int_{x \in \Omega_X} \mathbb{P}(A|X = x) f_X(x) dx$$

- **Continuous Law of Total Expectation:**

$$\mathbb{E}[X] = \int_{y \in \Omega_Y} \mathbb{E}[X|Y = y] f_Y(y) dy$$

- **Markov's Inequality:** Let X be a non-negative random variable, and $\alpha > 0$. Then,

$$\mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}$$

- **Chebyshev's Inequality:** Suppose Y is a random variable with $\mathbb{E}[Y] = \mu$ and $\text{Var}(Y) = \sigma^2$. Then, for any $\alpha > 0$,

$$\mathbb{P}(|Y - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}$$

- **(Multiplicative) Chernoff Bound:** Let $X_1, X_2, \dots, X_n$ be *independent* Bernoulli random variables.

Let $X = \sum_{i=1}^n X_i$, and $\mu = \mathbb{E}[X]$. Then, for any $0 \leq \delta \leq 1$,

- $\mathbb{P}(X \geq (1 + \delta)\mu) \leq e^{-\frac{\delta^2 \mu}{3}}$
- $\mathbb{P}(X \leq (1 - \delta)\mu) \leq e^{-\frac{\delta^2 \mu}{2}}$

1. Joint PMF's

Suppose X and Y have the following joint PMF:

X/Y	1	2	3
0	0	0.2	0.1
1	0.3	0	0.4

- (a) Identify the range of X (Ω_X), the range of Y (Ω_Y), and their joint range ($\Omega_{X,Y}$). **Solution:**

$$\Omega_X = \{0, 1\}, \Omega_Y = \{1, 2, 3\}, \text{ and } \Omega_{X,Y} = \{(0, 2), (0, 3), (1, 1), (1, 3)\}$$

- (b) Find the marginal PMF for X , $p_X(x)$ for $x \in \Omega_X$. **Solution:**

$$p_X(0) = \sum_y p_{X,Y}(0, y) = 0 + 0.2 + 0.1 = 0.3$$

$$p_X(1) = 1 - p_X(0) = 0.7$$

- (c) Find the marginal PMF for Y , $p_Y(y)$ for $y \in \Omega_Y$. **Solution:**

$$p_Y(1) = \sum_x p_{X,Y}(x, 1) = 0 + 0.3 = 0.3$$

$$p_Y(2) = \sum_x p_{X,Y}(x, 2) = 0.2 + 0 = 0.2$$

$$p_Y(3) = \sum_x p_{X,Y}(x, 3) = 0.1 + 0.4 = 0.5$$

- (d) Are X and Y independent? Why or why not? **Solution:**

No, since a necessary condition is that $\Omega_{X,Y} = \Omega_X \times \Omega_Y$.

- (e) Find $\mathbb{E}[X^3 Y]$. **Solution:**

Note that $X^3 = X$ since X takes values in $\{0, 1\}$.

$$\mathbb{E}[X^3Y] = \mathbb{E}[XY] = \sum_{(x,y) \in \Omega_{X,Y}} xy p_{X,Y}(x,y) = 1 \cdot 1 \cdot 0.3 + 1 \cdot 3 \cdot 0.4 = 1.5$$

2. Tail bounds

Suppose $X \sim \text{Binomial}(6, 0.4)$. We will bound $\mathbb{P}(X \geq 4)$ using the tail bounds we've learned, and compare this to the true result.

- (a) Give an upper bound for this probability using Markov's inequality. Why can we use Markov's inequality?

Solution:

We know that the expected value of a binomial distribution is np , so: $\mathbb{P}(X \geq 4) \leq \frac{\mathbb{E}[X]}{4} = \frac{2.4}{4} = 0.6$. We can use it since X is nonnegative.

- (b) Give an upper bound for this probability using Chebyshev's inequality. You may have to rearrange algebraically and it may result in a weaker bound. **Solution:**

$\mathbb{P}(X \geq 4) = \mathbb{P}(X - 2.4 \geq 1.6) \leq \mathbb{P}(|X - 2.4| \geq 1.6)$ we can add those absolute value signs because that only adds more possible values, so it is an upper bound on the probability of $X - 2.4 \geq 1.6$. Then, using Chebyshev's inequality we get:
 $\mathbb{P}(|X - 2.4| \geq 1.6) \leq \frac{\text{Var}(X)}{1.6^2} = \frac{1.44}{1.6^2} = 0.5625$

- (c) Give an upper bound for this probability using the Chernoff bound. **Solution:**

$$\mathbb{P}(X \geq 4) = \mathbb{P}(X \geq (1 + \frac{2}{3})2.4) \leq e^{-(\frac{2}{3})^2 \mathbb{E}[X]/3} = e^{-4 \times 2.4/27} \approx 0.7$$

- (d) Give the exact probability. **Solution:**

Since X is a binomial, we know it has a range from 0 to n (or in this case 0 to 6). Thus, the possible values to satisfy $X \geq 4$ are 4, 5, or 6. We plug in the PMF for each to get: $\mathbb{P}(X \geq 4) = \mathbb{P}(X = 4) + \mathbb{P}(X = 5) + \mathbb{P}(X = 6) = \binom{6}{4}(0.4)^4(0.6)^2 + \binom{6}{5}(0.4)^5(0.6) + \binom{6}{6}0.4^6 \approx 0.1792$

3. Exponential Tail Bounds

Let $X \sim \text{Exp}(\lambda)$ and $k > 1/\lambda$.

- (a) Use Markov's inequality to bound $\mathbb{P}(X \geq k)$.

Solution:

$$\mathbb{P}(X \geq k) \leq \frac{1}{\lambda k}$$

- (b) Use Chebyshev's inequality to bound $\mathbb{P}(X \geq k)$.

Solution:

$$\mathbb{P}(X \geq k) = \mathbb{P}\left(X - \frac{1}{\lambda} \geq k - \frac{1}{\lambda}\right) \leq \mathbb{P}\left(\left|X - \frac{1}{\lambda}\right| \geq k - \frac{1}{\lambda}\right) \leq \frac{1}{\lambda^2(k - 1/\lambda)^2} = \frac{1}{(\lambda k - 1)^2}$$

(c) What is the exact formula for $\mathbb{P}(X \geq k)$?

Solution:

$$\mathbb{P}(X \geq k) = e^{-\lambda k}$$

(d) For $\lambda k \geq 3$, how do the bounds given in parts (a), (b), and (c) compare?

Solution:

$$e^{-\lambda k} < \frac{1}{(\lambda k - 1)^2} < \frac{1}{\lambda k}$$

so Markov's inequality gives the worst bound.

4. Trinomial Distribution

A generalization of the Binomial model is when there is a sequence of n independent trials, but with three outcomes, where $\mathbb{P}(\text{outcome } i) = p_i$ for $i = 1, 2, 3$ and of course $p_1 + p_2 + p_3 = 1$. Let X_i be the number of times outcome i occurred for $i = 1, 2, 3$, where $X_1 + X_2 + X_3 = n$. Find the joint PMF $p_{X_1, X_2, X_3}(x_1, x_2, x_3)$ and specify its value for all $x_1, x_2, x_3 \in \mathbb{R}$. **Solution:**

Same argument as for the binomial PMF:

$$p_{X_1, X_2, X_3}(x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3} \prod_{i=1}^3 p_i^{x_i} = \frac{n!}{x_1! x_2! x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

where $x_1 + x_2 + x_3 = n$ and are nonnegative integers.

5. Do You “Urn” to Learn More About Probability?

Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let $X_i = 1$ if the i -th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

(a) X_1, X_2 **Solution:**

Here is one way of defining the joint pmf of X_1, X_2

$$\mathbb{P}(X_1 = 1, X_2 = 1) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 1 | X_1 = 1) = \frac{5}{13} \cdot \frac{4}{12} = \frac{20}{156}$$

$$\mathbb{P}(X_1 = 1, X_2 = 0) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 0 | X_1 = 1) = \frac{5}{13} \cdot \frac{8}{12} = \frac{40}{156}$$

$$\mathbb{P}(X_1 = 0, X_2 = 1) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 1 | X_1 = 0) = \frac{8}{13} \cdot \frac{5}{12} = \frac{40}{156}$$

$$\mathbb{P}(X_1 = 0, X_2 = 0) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0 | X_1 = 0) = \frac{8}{13} \cdot \frac{7}{12} = \frac{56}{156}$$

(b) X_1, X_2, X_3 **Solution:**

Instead of listing out all the individual probabilities, we could write a more compact formula for the pmf. In this problem, the denominator is always $P(13, k)$, where k is the number of random variables in the joint pmf. And the numerator is $P(5, i)$ times $P(8, j)$ where i and j are the number of 1s and 0s, respectively.

If we wish to compute $p_{X_1, X_2, X_3}(x_1, x_2, x_3)$, then the number of 1s (i.e., white balls) is $x_1 + x_2 + x_3$, and the number of 0s (i.e., red balls) is $(1 - x_1) + (1 - x_2) + (1 - x_3)$. Then, we can write the pmf as follows:

$$p_{X_1, X_2, X_3}(x_1, x_2, x_3) = \frac{10!}{13!} \cdot \frac{5!}{(5 - x_1 - x_2 - x_3)!} \cdot \frac{8!}{(5 + x_1 + x_2 + x_3)!}$$

6. Successes

Consider a sequence of independent Bernoulli trials, each of which is a success with probability p . Let X_1 be the number of failures preceding the first success, and let X_2 be the number of failures between the first 2 successes. Find the joint pmf of X_1 and X_2 . Write an expression for $E[\sqrt{X_1 X_2}]$. You can leave your answer in the form of a sum. **Solution:**

X_1 and X_2 take on two particular values x_1 and x_2 , when there are x_1 failures followed by one success, and then x_2 failures followed by one success. Since the Bernoulli trials are independent the joint pmf is

$$p_{X_1, X_2}(x_1, x_2) = (1 - p)^{x_1} p \cdot (1 - p)^{x_2} p = (1 - p)^{x_1 + x_2} p^2$$

for $(x_1, x_2) \in \Omega_{X_1, X_2} = \{0, 1, 2, \dots\} \times \{0, 1, 2, \dots\}$. By the definition of expectation

$$E[\sqrt{X_1 X_2}] = \sum_{(x_1, x_2) \in \Omega_{X_1, X_2}} \sqrt{x_1 x_2} \cdot (1 - p)^{x_1 + x_2} p^2.$$

7. Continuous joint density

The joint density of X and Y is given by

$$f_{X, Y}(x, y) = \begin{cases} x e^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the joint density of W and V is given by

$$f_{W, V}(w, v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent? Are W and V independent?

Solution:

For two random variables X, Y to be independent, we must have $f_{X, Y}(x, y) = f_X(x) f_Y(y)$ for all $x \in \Omega_X, y \in \Omega_Y$. Let's start with X and Y by finding their marginal PDFs. By definition, and using the fact that the joint PDF is 0 outside of $y > 0$, we get:

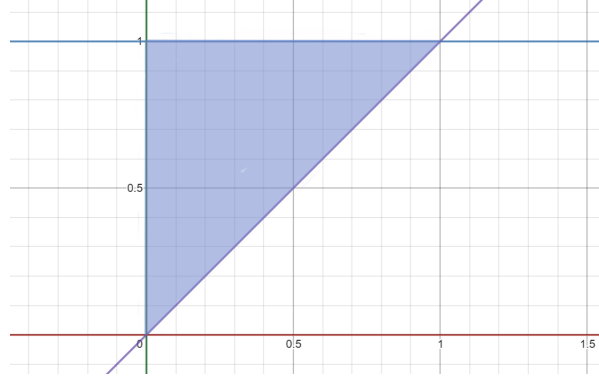
$$f_X(x) = \int_0^\infty x e^{-(x+y)} dy = e^{-x} x$$

We do the same to get the PDF of Y , again over the range $x > 0$:

$$f_Y(y) = \int_0^\infty x e^{-(x+y)} dx = e^{-y}$$

Since $e^{-x}x \cdot e^{-y} = xe^{-x-y} = xe^{-(x+y)}$ for all $x, y > 0$, X and Y are independent.

We can see that W and V are not independent simply by observing that $\Omega_W = (0, 1)$ and $\Omega_V = (0, 1)$, but $\Omega_{W,V}$ is not equal to their Cartesian product. Specifically, looking at their range of $f_{W,V}(w, v)$. Graphing it with w as the "x-axis" and v as the "y-axis", we see that :



The shaded area is where the joint pdf is strictly positive. Looking at it, we can see that it is not rectangular, and therefore it is not the case that $\Omega_{W,V} = \Omega_W \times \Omega_V$. Remember, the joint range being the Cartesian product of the marginal ranges is not sufficient for independence, but it is *necessary*. Therefore, this is enough to show that they are not independent.

8. Trapped Miner

A miner is trapped in a mine containing 3 doors.

- D_1 : The 1st door leads to a tunnel that will take him to safety after 3 hours.
- D_2 : The 2nd door leads to a tunnel that returns him to the mine after 5 hours.
- D_3 : The 3rd door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters $(12, \frac{1}{3})$.

At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety? **Solution:**

Let T = number of hours for the miner to reach safety. (T is a random variable)

Let D_i be the event the i^{th} door is chosen. $i \in \{1, 2, 3\}$. Finally, let T_3 be the time it takes to return to the mine in the third case only (a random variable). Note that the expectation of T_3 is $12 * \frac{1}{3}$ because it is binomially distributed with parameters $n = 12, p = \frac{1}{3}$. By Law of Total Expectation, linearity of expectation, and by applying the conditional expectations given by the problem statement:

$$\begin{aligned} \mathbb{E}[T] &= \mathbb{E}[T|D_1] \mathbb{P}(D_1) + \mathbb{E}[T|D_2] \mathbb{P}(D_2) + \mathbb{E}[T|D_3] \mathbb{P}(D_3) \\ &= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3 + T]) \cdot \frac{1}{3} \\ &= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3] + \mathbb{E}[T]) \cdot \frac{1}{3} \\ &= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (4 + \mathbb{E}[T]) \cdot \frac{1}{3} \end{aligned}$$

Solving this equation for $\mathbb{E}[T]$, we get

$$\mathbb{E}[T] = 12$$

Therefore, the expected number of hours for this miner to reach safety is 12.

9. Lemonade Stand

Suppose I run a lemonade stand, which costs me \$100 a day to operate. I sell a drink of lemonade for \$20. Every person who walks by my stand either buys a drink or doesn't (no one buys more than one). If it is raining, n_1 people walk by my stand, and each buys a drink independently with probability p_1 . If it isn't raining, n_2 people walk by my stand, and each buys a drink independently with probability p_2 . It rains each day with probability p_3 , independently of every other day. Let X be my profit over the next week. In terms of n_1, n_2, p_1, p_2 and p_3 , what is $\mathbb{E}[X]$? **Solution:**

Let R be the event it rains. Let X_i be how many drinks I sell on day i for $i = 1, \dots, 7$. We are interested in $X = \sum_{i=1}^7 (20X_i - 100)$. We have $X_i|R \sim \text{Binomial}(n_1, p_1)$, so $\mathbb{E}[X_i|R] = n_1 p_1$. Similarly, $X_i|R^C \sim \text{Binomial}(n_2, p_2)$, so $\mathbb{E}[X_i|R^C] = n_2 p_2$. By the law of total expectation,

$$\mu = \mathbb{E}[X_i] = \mathbb{E}[X_i|R] \mathbb{P}(R) + \mathbb{E}[X_i|R^C] \mathbb{P}(R^C) = n_1 p_1 p_3 + n_2 p_2 (1 - p_3)$$

Hence, by linearity of expectation,

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[\sum_{i=1}^7 (20X_i - 100)\right] = 20 \sum_{i=1}^7 \mathbb{E}[X_i] - 700 = 140\mu - 700 \\ &= 140 \cdot (n_1 p_1 p_3 + n_2 p_2 (1 - p_3)) - 700. \end{aligned}$$

10. 3 points on a line

Three points X_1, X_2, X_3 are selected at random on a line L (continuous independent uniform distributions). What is the probability that X_2 lies between X_1 and X_3 ? **Solution:**

Let $X_1, X_2, X_3 \sim \text{Unif}(0, 1)$.

$$\begin{aligned} \mathbb{P}(X_1 < X_2 < X_3) &= \int_{-\infty}^{\infty} \mathbb{P}(X_1 < X_2 < X_3 \mid X_2 = x) f_{X_2}(x) dx && \text{Continuous LoTP} \\ &= \int_{-\infty}^{\infty} \mathbb{P}(X_1 < x, X_3 > x) f_{X_2}(x) dx && \text{Independence of } X_1, X_2, X_3 \\ &= \int_{-\infty}^{\infty} \mathbb{P}(X_1 < x) \mathbb{P}(x < X_3) f_{X_2}(x) dx && \text{Independence of } X_1, X_3 \\ &= \int_{-\infty}^{\infty} F_{X_1}(x) (1 - F_{X_3}(x)) f_{X_2}(x) dx \\ &= \int_0^1 x (1 - x) 1 dx \\ &= \left. \frac{x^2}{2} - \frac{x^3}{3} \right|_0^1 = \frac{1}{6} \end{aligned}$$

11. Another Urn Question

An urn has 12 balls, 5 red ones and 7 green ones. Draw 3 balls. Let X denote the number of red balls in the sample. Compute $Var(X)$ when sampling is done:

- (a) With replacement
- (b) Without replacement

Solution:

(a) We start by introducing the indicator variables:

$$X_1 = \begin{cases} 1, & \text{first ball is red} \\ 0, & \text{first ball is green} \end{cases}$$

$$X_2 = \begin{cases} 1, & \text{second ball is red} \\ 0, & \text{second ball is green} \end{cases}$$

$$X_3 = \begin{cases} 1, & \text{third ball is red} \\ 0, & \text{third ball is green} \end{cases}$$

Then, note that X_1, X_2, X_3 are all $Ber(\frac{5}{12})$ and are independent.

$$Var(X_1) = Var(X_2) = Var(X_3) = p(1-p) = \frac{5}{12}(1 - \frac{5}{12}) = \frac{35}{144}$$

So:

$$\begin{aligned} Var(X) &= Var(X_1 + X_2 + X_3) \\ &= Var(X_1) + Var(X_2) + Var(X_3) [\text{since they are all independent}] \\ &= \frac{35}{144} + \frac{35}{144} + \frac{35}{144} = \frac{35}{48} \end{aligned}$$

- (b) We can consider this as taking out 3 balls. We will define X_1 for the first ball being red, X_2 for the second ball being red, and X_3 for the third ball being red. The independence of X_1, X_2 , and X_3 is no longer true. So, we will want to solve for the covariance matrix and find the sum of the entries. However, the marginal distributions of X_1, X_2 , and X_3 are $Ber(\frac{5}{12})$, since each has a probability $\frac{5}{12}$ of being red.

We can prove this because there are 5 red balls and 12 total balls. We want to calculate the probability that the i th ball is red, after we choose 3 balls from the urn. There are a total of $\mathbb{P}(12, 3)$ ways to order the 3 balls we picked from the urn. There are 5 ways to fix the red ball at the i th position. There are $\mathbb{P}(12 - 1, 3 - 1)$ ways to order the remaining 11 balls for the other 2 positions. This leaves us with:

$$\begin{aligned} \mathbb{P}(X_i = 1) &= \frac{5P(11, 2)}{P(12, 3)} \\ &= \frac{5 \frac{11!}{9!}}{\frac{12!}{9!}} \\ &= \frac{5}{12} \end{aligned}$$

Which means that $X_i \sim Ber(\frac{5}{12})$.

We calculated the variance for this above, and we have:

$$\text{Cov}(X_i, X_i) = \text{Var}(X_i) = \frac{35}{144}$$

Now, we have $\text{Cov}(X_1, X_2) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]$.

Note that $X_1 \cdot X_2$ is only 1 when both are 1. We start with 5 red balls of the 12 in the first choice, and then 4 of the remaining 11 in the second choice. So:

$$\mathbb{E}[X_1 X_2] = 1 \cdot \frac{5}{12} \cdot \frac{4}{11} = \frac{5}{33}$$

Then, since X_1 is the first choice, with 5 red balls of the 12:

$$\mathbb{E}[X_1] = \frac{5}{12}$$

For the second choice we have:

$$\begin{aligned} \mathbb{E}[X_2] &= \mathbb{E}[X_2 | X_1 = 1] \mathbb{P}(X_1 = 1) + \mathbb{E}[X_2 | X_1 = 0] \mathbb{P}(X_1 = 0) \\ &= \frac{4}{11} \cdot \frac{5}{12} + \frac{5}{11} \cdot \frac{7}{12} \\ &= \frac{5}{12} \end{aligned}$$

All together this means:

$$\begin{aligned} \text{Cov}(X_1, X_2) &= \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2] \\ &= -\frac{35}{1584} \end{aligned}$$

In fact, you will find that for any i , $\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \frac{5}{12}$, when we consider each of these variables marginally. Further, for any $i \neq j$, $\mathbb{E}[X_i X_j] = \frac{5}{33}$, we can solve for these manually considering all cases, or consider the similarity to the hat check or cat and mitten problem. We have:

$$\begin{aligned} \mathbb{E}[X_i X_j] &= \mathbb{P}(X_i = 1) \cdot \mathbb{P}(X_j = 1 | X_i = 1) \\ &= \frac{5}{12} \cdot \frac{4}{11} \\ &= \frac{5}{33} \end{aligned}$$

So, for any $i \neq j$, $\text{Cov}(X_i, X_j) = -\frac{35}{1584}$.

This gives us the following covariance matrix:

	X_1	X_2	X_3
X_1	$\frac{35}{144}$	$-\frac{35}{1584}$	$-\frac{35}{1584}$
X_2	$-\frac{35}{1584}$	$\frac{35}{144}$	$-\frac{35}{1584}$
X_3	$-\frac{35}{1584}$	$-\frac{35}{1584}$	$\frac{35}{144}$

So, we have the following:

$$\begin{aligned} Cov(X, X) &= Cov\left(\sum_{i=1}^3 X_i, \sum_{j=1}^3 X_j\right) \\ &= \sum_{i=1}^3 \sum_{j=1}^3 Cov(X_i, X_j) \\ &= 3 \cdot \frac{35}{144} + 2 \binom{3}{2} \cdot \left(-\frac{35}{1584}\right) \\ &= \frac{105}{176} \end{aligned}$$